Descriptive Finding

Using household death questions from surveys to assess adult mortality in periods of health crisis: An application for Peru, 2018–2022

José H.C. Monteiro da Silva
Helena Cruz Castanheira

This publication is part of the Special Collection on “Innovations in measuring adult mortality in countries with incomplete civil registration,” organized by Guest Editors Ashira Menashe Oren, Nandita Saikia, and Stéphane Helleringer.

© 2024 José H.C. Monteiro da Silva & Helena Cruz Castanheira.

This open-access work is published under the terms of the Creative Commons Attribution 3.0 Germany (CC BY 3.0 DE), which permits use, reproduction, and distribution in any medium, provided the original author(s) and source are given credit.
See https://creativecommons.org/licenses/by/3.0/de/legalcode.
## Contents

1 Introduction 216

2 Data 217
2.1 Household deaths 217
2.2 Sibling histories 218
2.3 218

3 Methods 219
3.1 Direct estimation from household deaths in a given period 219
3.2 Sibling survival method 220

4 Findings 220

5 Discussion 223

6 Acknowledgments 224

References 226
Using household death questions from surveys to assess adult mortality in periods of health crisis: An application for Peru, 2018–2022

José H.C. Monteiro da Silva
Helena Cruz Castanheira

Abstract

BACKGROUND
The COVID-19 pandemic highlighted the importance of civil registration and vital statistics (CRVS) systems for tracking and monitoring mortality outcomes during a public health crisis, especially in low- and middle-income countries. Alternative mortality data sources, such as censuses and surveys, offer an opportunity to assess the impact of health crises on countries with incomplete CRVS systems.

OBJECTIVE
Our aim is to show that data on retrospective household deaths collected in household surveys produce informative adult mortality numbers that can be useful in estimating mortality in the context of the COVID-19 pandemic in countries with incomplete CRVS systems, such as Peru.

METHODS
Using data on household deaths in the previous five years from the National Demographic and Family Health Surveys of Peru (ENDES) from 2018, 2019, 2021, and 2022, we estimate the probability of dying between ages 20 and 64 (\(45q_{20}\)) and compare the results with estimates from the World Population Prospects (WPP) 2022 revision and with estimates attained using the sibling survival method.

RESULTS
We verify that \(45q_{20}\) estimates from ENDES household death information fall close to those reported by the WPP 2022. However, these estimates have high confidence intervals due to the small sample size. The sibling survival method consistently estimates lower adult mortality probabilities, even in pandemic years. Despite the difference in

---

2 Latin American and Caribbean Demographic Centre, Population Division of the United Nations Economic Commission for Latin America and the Caribbean, Santiago, Chile. The views expressed in this paper are those of the author and do not necessarily reflect the views of the United Nations.
magnitude between the WPP 2022 estimates and those from household deaths, both provide a picture of an increase in the probability of dying among adults during the pandemic period. This is not reflected in estimates made using the sibling survival method.

CONTRIBUTION
Despite small sample size and irregularities in age-specific estimates, our work shows that survey questions on household deaths have a great potential for informing adult mortality over time in countries with deficient CRVS systems.

1. Introduction

Civil registration and vital statistics (CRVS) systems are the gold standard for the monitoring of several sustainable development goals. However, most low- and middle-income countries (LMICs), in particular those from sub-Saharan Africa and South Asia, still lag behind in the implementation of well-functioning CRVS systems for collecting basic information on the age and sex of the deceased (Mikkelsen et al. 2015). The COVID-19 pandemic highlighted the relevance of CRVS systems for tracking and monitoring mortality outcomes during a public health crisis, and the absence of properly functioning systems in LMICs made it difficult to retrieve numbers showing the direct and indirect impact of COVID-19 on mortality. Therefore alternative mortality data sources, such as censuses and surveys, offer an opportunity to assess the impact of health crises on these countries (Technical Advisory Group on COVID-19 Mortality Assessment 2022).

The United Nations recommends the inclusion of a questionnaire on household deaths in the previous 12 months for population and housing censuses to capture mortality information on countries that do not have “satisfactory continuous death statistics” from CRVS systems (United Nations 2017). Although the question is expected to involve some underreporting of deaths, such as that produced by design because information is only for households existing at the time of the inquiry (selection error) – deaths from unipersonal households that were dissolved during the recall period are not captured (United Nations 2004) – it offers multiple advantages for the study and monitoring of mortality at the national and subnational levels and for subgroups of the population. For instance, it allows for the direct estimation of mortality rates for all age groups using the numerator and denominator from the same data source, reducing potential biases that emerge from mismatches between sources. Also, these data offer the opportunity to work with lower levels of disaggregation (geographic and socioeconomic), since national censuses collect relevant household socioeconomic information, such as education and living conditions (Queiroz and Sacco 2018; Queiroz and Sawyer 2012). Questions on
household deaths in a recent period have been included in the most recent census rounds of several Latin American countries, and they have provided valuable insights for mortality estimation in the region (Queiroz and Sacco 2018).

Regarding household surveys, few studies use questions about deaths in the household in a specific period. Usually, demographic and health surveys (DHS) have basic questions regarding deaths of siblings, which are used to estimate mortality through the sibling survival method (Moultrie et al. 2013). This method is particularly useful in countries that do not have CRVS data and has been broadly used for adult mortality estimation at national and subnational levels in several LMICs (Masquelier 2013; Menashe-Oren and Masquelier 2022). Nevertheless, the mortality numbers from the sibling survival method are likely to be underestimated due to incorrect sibling age reporting and age of death reporting (Helleringer et al. 2014; Masquelier et al. 2021).

In this descriptive note, we want to show that, despite the data quality issues – due to the challenges of capturing deaths in small samples in household surveys and the underreporting of deaths at older ages and in single-person households (Helleringer 2022; Hill, Choi, and Timaeus 2005; Lankoandé et al. 2022; Timaeus 1991) – reported retrospective household deaths produce informative adult mortality estimates and can be particularly useful in estimating mortality in the context of a public health crisis such as the COVID-19 pandemic. We provide an application of the use of this information for Peru. Despite recent advances in establishing a functioning CRVS system, the relative death registration completeness of Peru lies between 70% and 80%, with high variability across regions (Castanheira and Monteiro da Silva 2022). The use of surveys can complement eventual gaps that emerge from the CRVS data.

Using data on household deaths in the previous five years from the National Demographic and Family Health Surveys of Peru (Encuesta Demográfica y de Salud Familiar in Spanish; ENDES) from 2018, 2019, 2021, and 2022, we estimate the probability of dying between ages 20 and 64 in five different periods and compare these results with estimates from the World Population Prospects (WPP) 2022 revision (United Nations 2022) and with estimates made using the sibling survival method.

2. Data

2.1 Household deaths

We used the question on household deaths over a specific period from the ENDES of Peru from 2018, 2019, 2021, and 2022. This specific inquiry on household deaths was first included in 2017, but the microdata are publicly available only for 2018 onward.³

The ENDES of 2018, 2019, 2020, 2021, and 2022 asked questions on the age and sex of the deceased and the month and year of the deaths that occurred in the household in the five years before the date of the interview. Table 1 shows some basic information about interview period, period of reported deaths, population, and death counts in the sample for the years 2018–2022.

The 2020 survey was excluded from our analysis due to data quality issues, especially because of the difficulties in collecting data during the first year of the pandemic, which were reported to the authors by members of the National Statistics Office of Peru. The main difference between the 2020 ENDES and the other ENDES is the data collection mode. Collection was done mainly by phone interview in 2020, affecting comparability across surveys. Moreover, the 2020 survey had higher non-response rates.

Table 1: Basic metadata of ENDES (survey year, period of interview, period of deaths, and sampled population and deaths), Peru, 2018–2022

<table>
<thead>
<tr>
<th>Survey year</th>
<th>Period of interview</th>
<th>Period of deaths</th>
<th>Sample size (N)</th>
<th>Deaths (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2022</td>
<td>Jan 2022 – Mar 2023</td>
<td>Jan 2017 – Dec 2022</td>
<td>139,618</td>
<td>2,931</td>
</tr>
</tbody>
</table>

2.2 Sibling histories

The mortality estimates made using household deaths over the previous five years were compared to estimates made using the sibling survival method. For that, we used sibling histories data collected by the same surveys (ENDES) on the maternal mortality module, using the same years (2018, 2019, 2021, and 2022). Following DHS protocol, the questionnaire collected information from women aged 12–49 about their siblings’ age, sex, survival status, and, in the case of sibling death, the date (year and month) of death.

2.3 World Population Prospects 2022 revision

Because of WPP 2022’s importance and widespread use in international comparisons, its mortality estimates (United Nations 2022) were used as a benchmark to assess the robustness of the household death and sibling survival estimation.

The Peruvian life tables in WPP 2022 were estimated using a combination of methods for child, adult, and old age mortality. The logistic-quadratic (LogQuad) relational life table approach (Wilmoth et al. 2012) was used with child and adult
probability of dying as inputs. The probabilities of dying under 5 years old for the years considered in this paper were estimated using the birth histories of the ENDES and the UN-IGME estimates (childmortality.org). The probabilities of dying between 15 and 59 years old were based on different data sources that were modeled using a Bayesian hierarchical model (Chao 2022) for each sex from 1950 to 2021 (United Nations 2022b).

3. Methods

3.1 Direct estimation from household deaths in a given period

Using information on deaths in the previous five years from these surveys, we estimated the mortality rates ($n_{m_x}$, age-specific mortality rate between ages $x$ and $x+n$) between ages 20 to 64 in five-year age groups for five periods of time: January 2017 to December 2018 (reference date January 2018, using the 2018 survey), January 2018 to December 2019 (reference date January 2019, using the 2019 survey), January 2019 to December 2020 (reference date January 2020, using the 2021 survey), January 2020 to December 2021 (reference date January 2021, using the 2021 survey), and January 2021 to December 2022 (reference date January 2022, using the 2022 survey). We chose the most recent two-year time span in each survey instead of one year to have a higher sample size and thus more robust estimates by sex and age groups. We did not use the five-year time span of each survey to better capture the pandemic effect in the 2021 and 2022 ENDES surveys and to avoid potential recall biases.

The point estimates of mortality rates were calculated by dividing the death counts within the period by the person-year exposures computed by following individuals’ lifelines (following the Lexis diagram) and considering the sample weights. Further, we estimated the 95% confidence intervals in a three-step approach. First, we estimated univariate Poisson models for the individual deaths for each sex and age group. Thus the exponentials of the intercepts from these models have the same values as the point estimates for the mortality rates we described, and they also provide the variance-covariance matrices accounting for the survey sampling design. In the second step, we simulated 10,000 values from a normal distribution with the mean equal to the exponential of the intercepts of step 1 and with variance from the models of step 1. We then used a third step to include the uncertainty brought by the stochasticity of mortality (Hendi 2023). Assuming that deaths follow a Poisson distribution, we ran 10,000 simulations using the person-year exposures and the mortality rates estimated in step 2. Finally, we retrieved the 95% confidence intervals from the estimates of step 3 by calculating the percentiles 2.5% and 97.5%.

Afterward, we converted these mortality rates into the probabilities of dying ($n_{q_x}$, probability of dying between ages $x$ and $x+n$) between ages 20 and 64 in five-year age groups ($n = 5$), assuming that life table deaths occurred on average in the middle of the
Then we calculated the probability of surviving from age $x$ to $x + n$ ($n p_x = 1 - n q_x$), and from the product of the survival probabilities between 20 and 64 we estimated the probability of dying between ages 20 and 64:

$$45 q_{20} = 1 - \prod_{x=20}^{60} n p_x$$

Finally, we used the 95% confidence intervals of mortality rates to estimate lower and upper bounds of the adult mortality probabilities.$^4$

### 3.2 Sibling survival method

With the sibling histories data, the estimation of adult mortality is straightforward and does not require any indirect method (Moultrie et al. 2013). Using the sibling histories, we can follow the lifelines of individuals in the Lexis diagram and count the number of deaths (numerator of mortality rates) and the population exposure (denominator of mortality rates) over a specified period. Once we have the mortality rates, the estimation of the adult probability of dying can be performed using Equations (1) and (2). For this exercise, we used the R package `demogsurv` for estimating adult mortality using the sibling survival method. We also estimated 95% confidence intervals using the `jackknife` strategy. We compared our estimates of $45 q_{20}$ with those estimated using the sibling histories and with those estimated by WPP 2022.

### 4. Findings

Figure 1 shows the estimated mortality rates using the two-year window for each ENDES survey round (2018, 2019, 2021, and 2022). We verify that the mortality rates estimated with the two-year window are irregular. However, the estimates from pandemic years (2020, 2021, and 2022) are consistently above the other curves, especially among the older age-groups (45+). The mortality rates for females below age 35 seem to be more irregular than for older ages.

---

$^4$ The lower and upper bounds of the adult mortality probabilities cannot be interpreted as 95% confidence intervals.
Figure 1: Age-specific mortality rates (five-year age groups between 20 and 64) – point estimates by sex from the household death inquiry – Peru, ENDES, 2018, 2019, 2021, and 2022

Figure 2 reports the estimated adult mortality probabilities using information from sibling histories, the WPP 2022 estimates, and household death data (using the mortality rates of Figure 1). The intervals between the lower and upper bounds for the estimates from household deaths are relatively large due to the uncertainty brought both by the sample design (and its small size) and the stochasticity of mortality itself. The point estimates of the ENDES surveys, however, are relatively close to the WPP 2022 estimates. The estimates using sibling histories, on the other hand, consistently report very low adult mortality probabilities.
We then compare the ratios between $45q_{20}$ estimates from 2020 and 2021 with those from 2019 for sibling histories, WPP 2022, and household death estimates to get a picture of the pandemic effect on Peru’s adult mortality probability (Table 2). We notice that the ratios are higher for estimates from household deaths. Despite the difference in magnitude between these sources, WPP 2022 and household death estimates both provide a picture of an increase in the probability of dying among adults during the pandemic period. For males, the probability of dying increased by around 40% from 2019 to 2021 according to WPP 2022 and by around 145% from the estimates of household deaths. For females, these probabilities increased by 50% in WPP 2022 and by 123% as estimated using household deaths from ENDES 2022 and 2021 surveys. The sibling history estimates, however, did not seem to capture the pandemic effect, except for females in 2022.

<table>
<thead>
<tr>
<th>Source</th>
<th>Females</th>
<th></th>
<th></th>
<th>Males</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2020</td>
<td>2021</td>
<td>2022</td>
<td>2020</td>
<td>2021</td>
<td>2022</td>
</tr>
<tr>
<td>Sibling histories</td>
<td>0.60</td>
<td>0.67</td>
<td>1.32</td>
<td>0.58</td>
<td>1.02</td>
<td>0.75</td>
</tr>
<tr>
<td>WPP 2022</td>
<td>1.29</td>
<td>1.67</td>
<td>1.50</td>
<td>1.45</td>
<td>1.57</td>
<td>1.40</td>
</tr>
<tr>
<td>Household deaths</td>
<td>2.10</td>
<td>2.01</td>
<td>2.22</td>
<td>2.83</td>
<td>2.77</td>
<td>2.44</td>
</tr>
</tbody>
</table>

5. Discussion

This research note has shown that the use of household death questions over a specific period is a valuable tool for estimating adult mortality in Peru. Using the example of Peru, we estimated the probability of dying between 20 and 64 using the ENDES surveys from 2018–2022 (2020 was excluded due to data collection issues during the pandemic, as reported by the National Statistics Office team) and compared our results with those estimated using information on sibling histories and those published by WPP 2022. In addition, we estimated the confidence intervals of age-specific mortality rates using the sample design of each ENDES survey and accounting for uncertainty of mortality (Hendi 2023).

We verified that the estimates of adult mortality probability using information on household deaths from the ENDES survey were similar to those reported by WPP 2022. This is likely because WPP 2022 derives its estimates from adjusted CRVS data – previous findings for Peru reported similar adult mortality levels using household death information and adjusted CRVS data for the period 2016–2018 (Castanheira and Monteio da Silva 2022). Estimates from household death information and from WPP 2022 were both higher than estimates from sibling histories. Previous evidence from record linkage studies have documented that adult mortality numbers from sibling histories tend to be underestimated due to recall biases and misreporting of siblings’ ages at death (Helleringer et al. 2014; Masquelier et al. 2021). To our knowledge, there are no record linkage studies that evaluate the quality and accuracy of household death mortality estimates. Lankoandé et al. (2022) performed an assessment using record linkage studies in Burkina Faso and showed that mortality numbers using household deaths from the 2006 census were underestimated.

Our results show that the information on household deaths collected by the ENDES surveys of 2021 and 2022 captured the pandemic effects on adult mortality in Peru. Peru was one of the most severely affected countries of Latin America during the COVID-19 pandemic – estimates suggest that the country observed close to 100% excess deaths between 2020 and 2021 (Msembali et al. 2023), values close to the ones we observed for the adult mortality probability in our calculations. Also, in accordance with previous
evidence, our household death estimates reported a higher increase in mortality risk for males than for females during the COVID-19 pandemic (Ramírez-Soto, Arroyo-Hernández, and Ortega-Cáceres 2021).

There are three main limitations regarding the use of information on household deaths for mortality estimation. First, death is a relatively rare vital event, which makes it hard to capture in household surveys with much smaller samples than those of censuses. Second, previous works in other countries and contexts have reported that household surveys and censuses consistently underestimate the level of mortality (Timæus 1991). Hence we encourage the estimation of confidence intervals whenever possible for capturing the uncertainty intrinsic to this type of data and to strengthen the reported estimates. Also, other demographic techniques can be applied for assessing the underestimation bias (Moultrie at al. 2013). Third, because of the above-mentioned factors, the age-specific mortality rates can be irregular and show unreliable fluctuations. One strategy to minimize the age fluctuation would be to increase the estimation period. However, in the case of our example, that would bias the estimates for the COVID-19 pandemic years. Another alternative would be to combine the survey information with indirect methods (such as model life tables) or to use the estimated age-specific rates within a mortality model (such as a Gompertz model) for generating smooth mortality curves.

Finally, we encourage the inclusion of household mortality inquiries in censuses and household surveys for mortality estimation. These questions have proven to be reliable and can provide important information at the national and subnational levels. In this work, we show that this information is also valuable in the context of a health crisis, such as the case of the COVID-19 pandemic. With the appropriate question, adequate training, and careful fieldwork, the question can provide valuable information on age-specific mortality rates by sex.

In comparison with other methods used to estimate adult mortality from censuses and surveys (United Nations 2004) and direct distribution methods used to evaluate the completeness of CRVS data (Castanheira and Monteiro da Silva 2022), the question on household deaths has a great potential for informing on adult mortality over time in countries with deficient CRVS systems.

6. Acknowledgments

We especially thank Eugenio Paglino and the two anonymous reviewers for comments that contributed to improving the paper. José HC Monteiro da Silva gratefully

---

5 Other methods available for estimating adult mortality rates: using (a) model life tables, (b) survival of parents, (c) survival of siblings.
acknowledges the resources provided by the International Max Planck Research School for Population, Health and Data Science (IMPRS-PHDS).
References


